

Technical Validation

Dell EMC VxRail and Intel Optane Persistent Memory

Enabling High Performance for HCI Workloads

By Kerry Dolan, Senior IT Validation Analyst

July 2020

This ESG Technical Validation was commissioned by Dell and Intel and is distributed under license from ESG.

Contents

Introduction	3
Background	3
Solution: Dell EMC VxRail Hyperconverged Infrastructure with Intel Optane PMem.....	4
Dell EMC VxRail	4
Intel Optane Persistent Memory	5
ESG Technical Validation.....	7
App Direct Mode Performance and Scalability Testing	7
Simulated Application Testing	11
The Bigger Truth.....	15
Appendix	16

ESG Technical Validations

The goal of ESG Technical Validations is to educate IT professionals about information technology solutions for companies of all types and sizes. ESG Technical Validations are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objectives are to explore some of the more valuable features and functions of IT solutions, show how they can be used to solve real customer problems, and identify any areas needing improvement. The ESG Validation Team’s expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.

Introduction

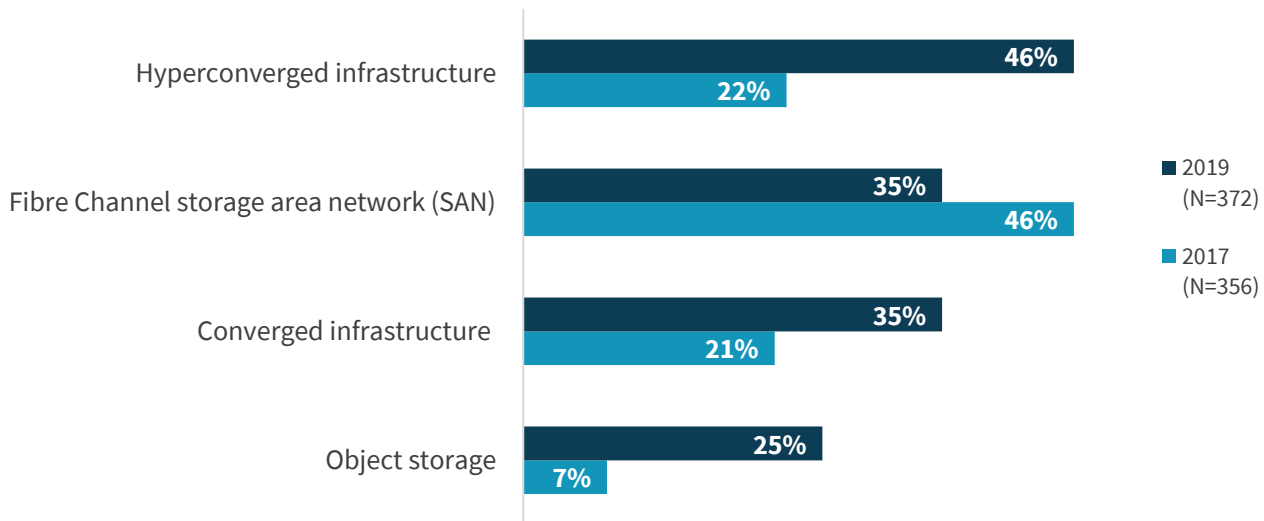
This report documents ESG’s validation of Dell Technologies performance testing of the VxRail Hyperconverged Infrastructure (HCI) configured with Intel Optane persistent memory (PMem). The high performance that PMem adds enables organizations to bring mission-critical, high-performance workloads to VxRail, increasing the opportunities to gain from VxRail’s management, efficiency, and flexibility benefits.

Background

Across all industries, organizations have recognized the benefits of HCI as they modernize data centers. The ability to consolidate multiple virtual workloads on a single platform brings agility and scalability to match the speed of business today. Instead of individually managed silos of infrastructure (compute, storage, and network) that IT must separately evaluate, procure, test, deploy, tune, and manage, HCI solutions like Dell EMC VxRail provide a centrally managed, single solution with software-defined resources that are easy to deploy, scale, and manage. These features also endow HCI solutions with a lower cost of ownership advantage. It is no surprise, then, that ESG research shows a considerable increase in HCI usage over the past few years. While in 2017 22% of respondents were employing HCI on-premises, the responses in 2019 had more than doubled to 46% (see Figure 1).¹ Organizations see HCI as a key contributor to digital transformation and are using it to power mission-critical production applications.

Figure 1. Increase in Use of Hyperconverged Infrastructures

On-premises data center storage technologies currently in use, 2017 vs. 2019.
(Percent of respondents)



Source: Enterprise Strategy Group

However, when multiple applications compete for resources, an “I/O blender effect” can occur, degrading performance. As a result, organizations have historically avoided placing on HCI applications that need consistent high performance, such as real-time analytics, online transaction processing (OLTP), content delivery, and high-performance computing (HPC). In order for these applications to run on HCI, a hyperconverged solution must offer high performance; this would expand the opportunity to gain from HCI’s efficiency, flexibility, and scalability.

¹ Source: ESG Research Report, [Data Storage Trends in an Increasingly Hybrid Cloud World](#), March 2020.

Solution: Dell EMC VxRail Hyperconverged Infrastructure with Intel Optane PMem

VxRail is an HCI solution that delivers agility, scalability, simplicity, and operational efficiency. Jointly developed by Dell EMC, VMware, and Intel, VxRail now offers configurations using Intel Optane persistent memory, increasing both performance and memory capacity while maintaining data persistence, at a significantly lower cost than DRAM.

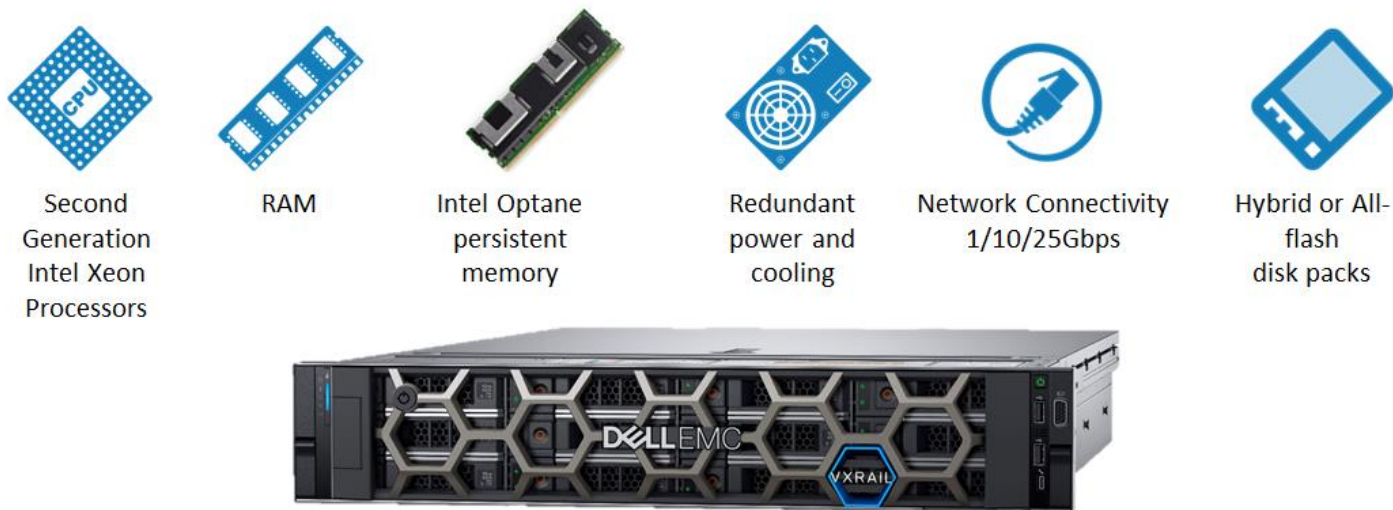
Dell EMC VxRail

VxRail provides turnkey infrastructure for a variety of modern, virtualized applications including databases, virtual desktop infrastructure (VDI), data analytics, and artificial intelligence (AI) at the core, edge, or cloud. It offers the fastest, simplest VMware environment, including management using the familiar vCenter interface. VxRail includes virtualized compute, networking, and storage; VxRail’s foundation is Dell PowerEdge servers with Second Generation Intel Xeon Scalable Processors and VxRail HCI System Software.

Two VxRail models (E560F/N and P570F/N) now offer configurations that include Intel Optane PMem, adding the performance necessary to support performance-intensive workloads such as OLTP, real-time analytics, in-memory databases, and HPC. Intel Optane PMem is available for VxRail in 128GB or 256GB configurations up to 6TiB, with 512GB cards available in the coming months. Two operational modes can be utilized: Memory Mode, which provides additional volatile memory that has a clear cost advantage over DRAM, and App Direct Mode, which can be used as low-latency, persistent storage.

Figure 2. Dell EMC VxRail Hyperconverged Infrastructure with Intel Optane PMem

Dell EMC VxRail



Using modular building blocks, organizations can non-disruptively scale capacity and performance from three to 64 nodes in a single cluster. VxRail includes the VxRail HCI System Software that keeps infrastructure in a continuously validated state. It provides intelligent lifecycle management, including automated, non-disruptive upgrades, patches, node additions, and node retirement. It also provides artificial intelligence- and machine learning (AI/ML)-driven health reporting, designed to help customers maintain optimal performance and availability. Options include:

- 4-112 Second Generation Intel Xeon Scalable Processor cores; single, dual, or quad socket.

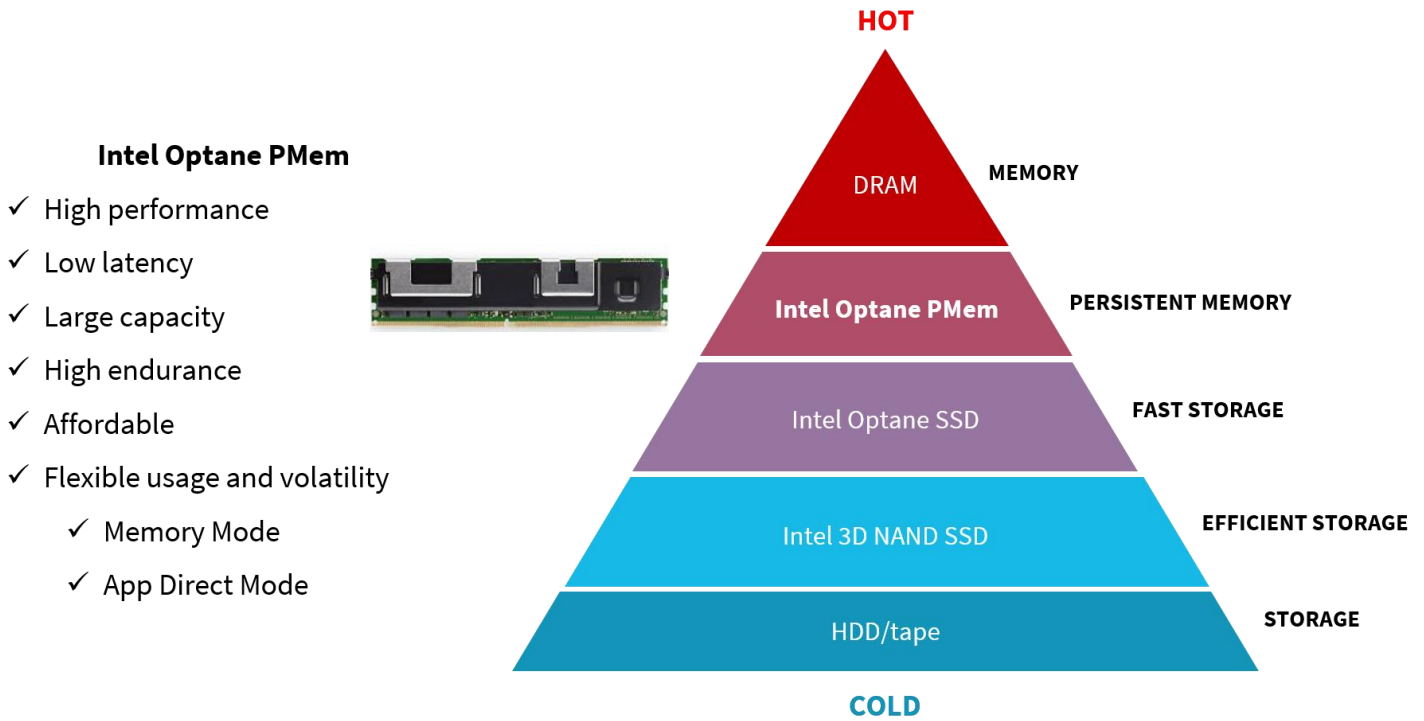
- 128GB or 256GB Intel Optane PMem cards, up to 6TiB.
- NVMe, SAS, or SATA drives, hybrid or all flash, up to 7.68 TB.
- Intel Optane SSDs.

Intel Optane Persistent Memory

At a high level, Intel’s Optane technology allows memory cells to be individually addressed in a dense, transistor-less, stackable design. Intel offers Intel Optane SSD and persistent memory, both providing faster storage than traditional NAND SSD. Intel Optane PMem sits between memory and storage, creating an additional performance tier between DRAM and SSD that can be non-volatile. Supported operating systems include MS Windows Server 2019, VMware ESXi version 6.7, RHEL, Ubuntu, Oracle, CentOS, and SLES Linux.

Intel Optane PMem devices are installed in the memory channel in the DDR4 DRAM form factor and can be combined with conventional DDR4 DIMMs. Each socket supports one to six devices, for up to 1.5 TiB per socket; the 512GB DIMM will offer 3 TiB per socket, double the maximum density of current DRAM DIMMs. Intel Optane PMem DIMMs also include 256-AES hardware encryption. Different from traditional DRAM, Intel Optane PMem offers both data persistence (so that data is retained during power loss or restart), and high density, with a lower cost per GB. It offers high performance, low latency, and better endurance for write-intensive workloads. By delivering more performance in a small footprint, Intel Optane PMem can improve total cost of ownership, by enabling higher workload consolidation rates and minimizing software licensing, node deployment, and power/cooling/data center costs.

Figure 3. Intel Optane Persistent Memory



Intel Optane PMem supports two addressing modes:

Benefits

Memory Mode provides expanded memory capacity, enabling organizations to run more VMs or increase memory per VM at a lower cost than DRAM. It also provides support for larger databases and greater VM, container, and application density. VDI, virtualized databases, and a big data analytics are good use cases for Memory Mode.

Memory Mode. This mode functions as an expanded, volatile system memory, enabling organizations to increase memory capacity in a denser footprint and at a lower cost than DRAM. It requires no changes to application software. In this mode, DRAM serves as a cache for the most frequently accessed data, while providing expanded main memory capacity. Cache

management is done by the Intel Xeon Scalable Processor memory controller. When data is requested, DRAM cache is checked first; if data is not there, it is read from Intel Optane PMem.

App Direct Mode. App Direct Mode provides data persistence

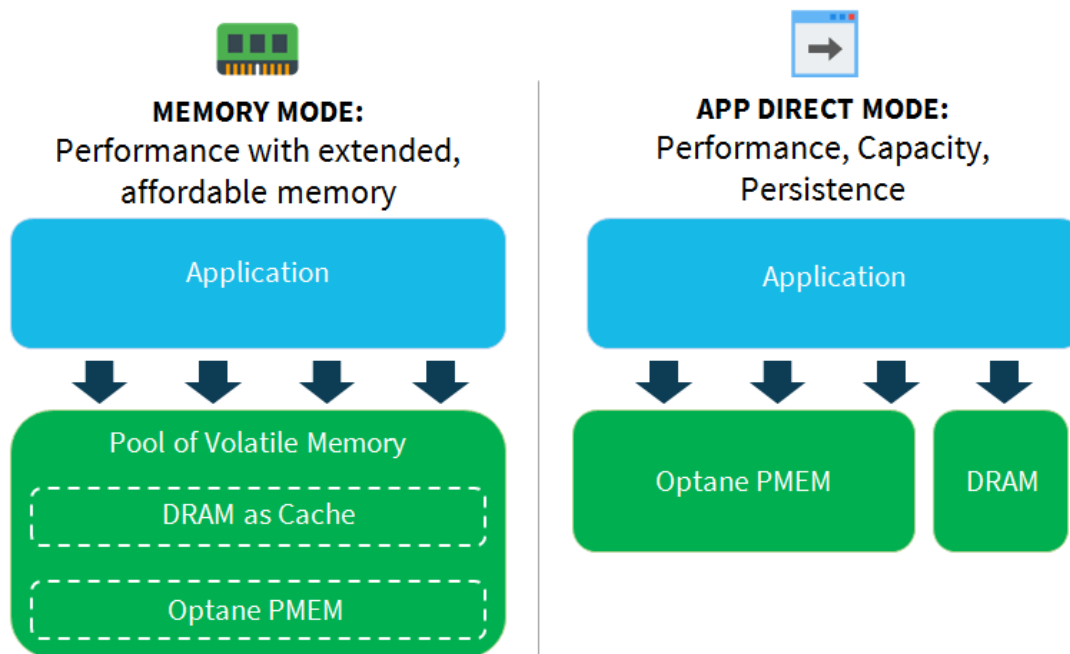
during a VM failure, eliminating the need to recover data during application restart. Applications are responsible for protecting data from a host failure using replication and other methods. In App Direct Mode, both Intel Optane PMem and DRAM are independent memory resources under direct load/store control of the application. It can be consumed as byte-addressable memory or as block storage and can be directly accessed by PMem-aware applications.

Because App Direct Mode expands the capacity of persistent memory, applications and operating systems recognize two types of direct-load memory—DRAM and PMem—and can send writes to the preferred one. Both are counted as total platform memory. Data that needs the lowest latency and does not require persistence can be sent to DRAM, while data that needs persistence and can support slightly higher latency can be directed to PMem. Applications that are optimized for Intel Optane PMem gain faster memory access with lower software I/O overhead. Byte-addressable App Direct Mode requires support for a persistent, memory-aware file system; supporting operating systems include vSphere 6.7U3, Windows Server 2019 NTFS DAX, Red Hat Enterprise Linux 7.6 or later, SuSe Enterprise Linux 12 SP4 or later, and Ubuntu 18.04 or later. Database support includes SAP HANA 2 SPS 4+, Redis, Aerospike, SQL Server 2019, and Apache Cassandra. VMware support includes vMotion and Distributed Resource Scheduler (DRS), with Virtual SAN (VSAN) support on the near-term roadmap.

Benefits

App Direct Mode brings large data sets closer to the CPU for faster processing and better insight, without the high cost of DRAM. Intel Optane PMem is more durable than NAND SSD for writes, and the native persistence delivers faster recovery and less downtime since it eliminates reloading databases into memory after shutdown. Use cases include in-memory databases, real-time analytics, OLTP, journaling, database log acceleration, HBASE hash tables, and high-performance computing.

Figure 4. Intel Optane PMem: Memory Mode and App Direct Mode



Configuring Intel Optane PMem

Intel Optane PMem is automatically recognized by the OS, and administrators configure the PMem operating mode using BIOS memory management tools. It can be designated as 100% App Direct Mode or 100% Memory Mode. Regions (defined as groups of PMem cards) can be configured to operate as interleaved (in which all PMem cards are seen as a single memory space with I/O striped across them) or non-interleaved (in which each PMem card is seen as a separate entity, with I/O filling one card before going to the next).

ESG Technical Validation

ESG audited results of performance testing by the VxRail team using Intel Optane PMem in App Direct Mode. Testing was designed to evaluate IOPS, throughput, scalability, and latency. *(Please note that an updated ESG report will be coming soon that includes Memory Mode performance testing.)*

First, we viewed a baseline of VxRail performance with all-flash disk groups using NVMe cache and SSDs. This was done using four VxRail P570Fs with two disk groups, each having 800GB NVMe cache and two 7.68TB SSD. Each VxRail host attained 1) 152K RR 4K IOPS and 40K RW 4K IOPS; and 2) 2.4 GB/s RR 64K throughput and 1.6 GB/s RW 64K throughput. These demonstrate impressive performance for an all-flash system running many workloads.

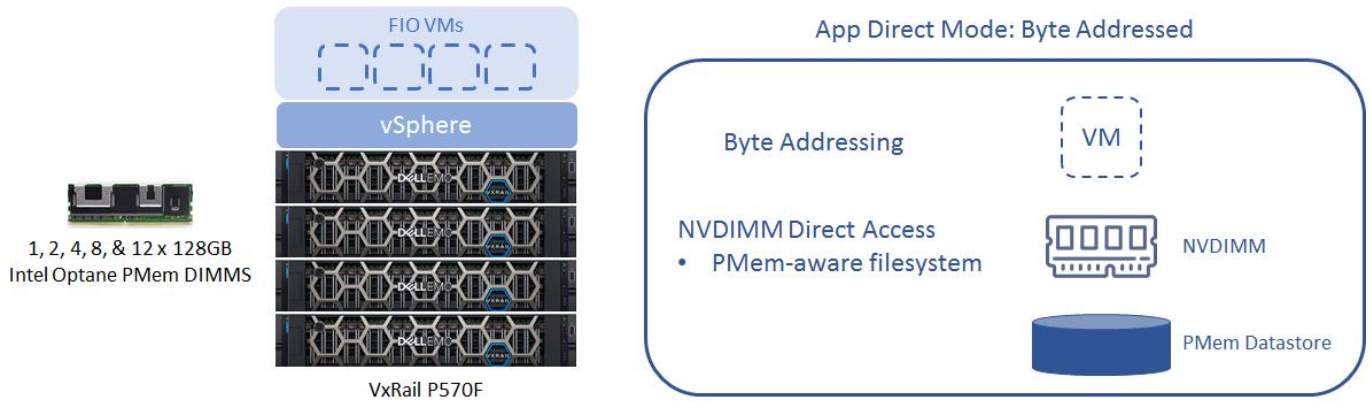
App Direct Mode Performance and Scalability Testing

Next, we viewed VxRail testing with Intel Optane PMem. The first set of tests was conducted using a four-node VxRail model P570F configured with dual Intel Xeon Gold 6254 CPUs, 800GB NVMe cache, and 23.8TB SAS SSD.² We evaluated the performance of the system in two ways: 1) using a single 128GB PMem card, and 2) scaling the number of 128GB PMem cards from one to 12. PMem was configured as byte-addressable NVDIMM using a PMem-aware DAX filesystem. The FIO test harness was used to generate workloads using one FIO VM per socket, with one disk per VM, and 80% working set size. I/O characterizations included 4KB and 64KB, 100% random read and 100% random write. It should be noted that

² See the Appendix for details of all test environments.

VxRail created and assigned non-uniform memory access (NUMA) nodes to tie together specific CPU and memory; in this testing, we confined each FIO NVDIMM to a single NUMA node. This ensures that VMs access PMem on a local socket for fast access.³

Figure 5. VxRail with Intel Optane PMem Test Bed



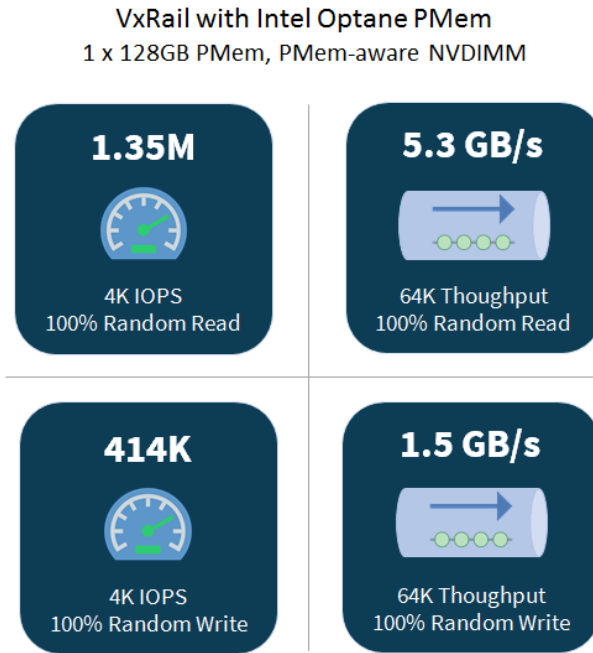
Source: Enterprise Strategy Group

We tested a single PMem card to find the “four corners” raw performance, a common assessment of basic horsepower of the system. As shown in Figure 6, the system supported high performance for both small block IOPS (more than 1.3 million read IOPS and more than 400K write IOPS) and larger block throughput (more than 5 GB/s for reads and more than 1.5 GB/s for writes).

The ability to support more than a million IOPS and multiple GB/s throughput with a single PMem card demonstrates that PMem can add significant performance to an HCI deployment at low cost.

³ For additional information regarding VMware NVDIMM configurations, please see <https://kb.vmware.com/s/article/78094>.

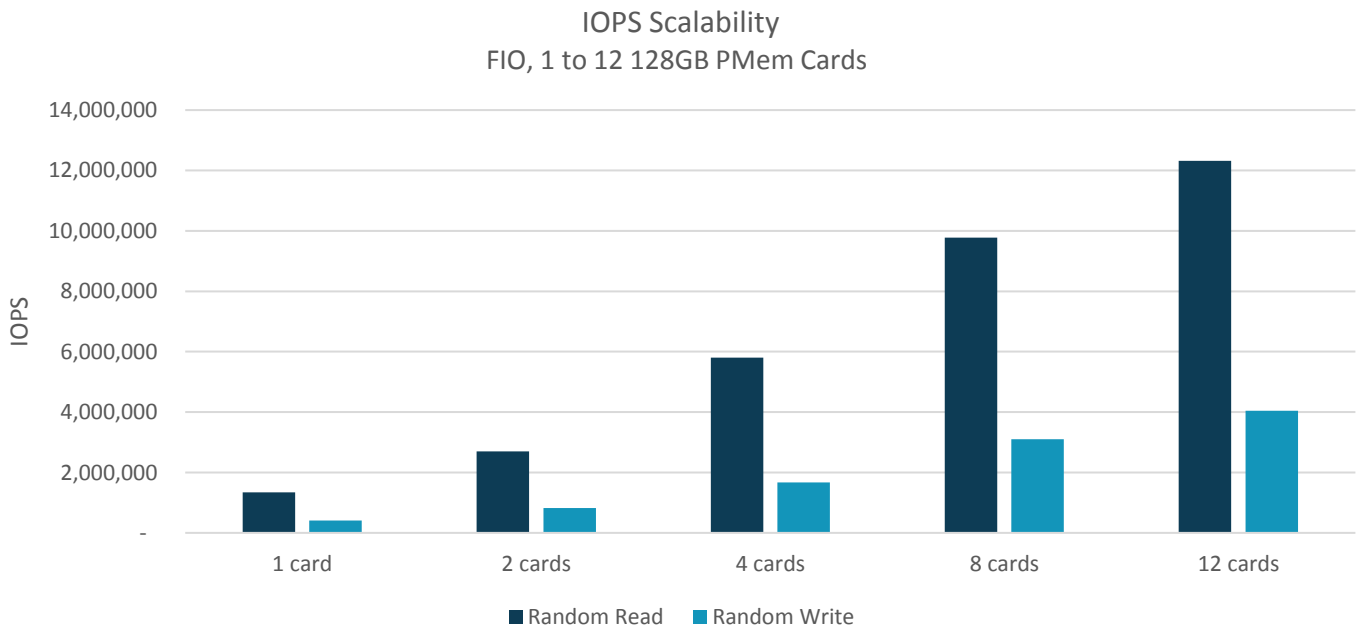
Figure 6. High-performance IOPS and Throughput, Single PMem Card



Source: Enterprise Strategy Group

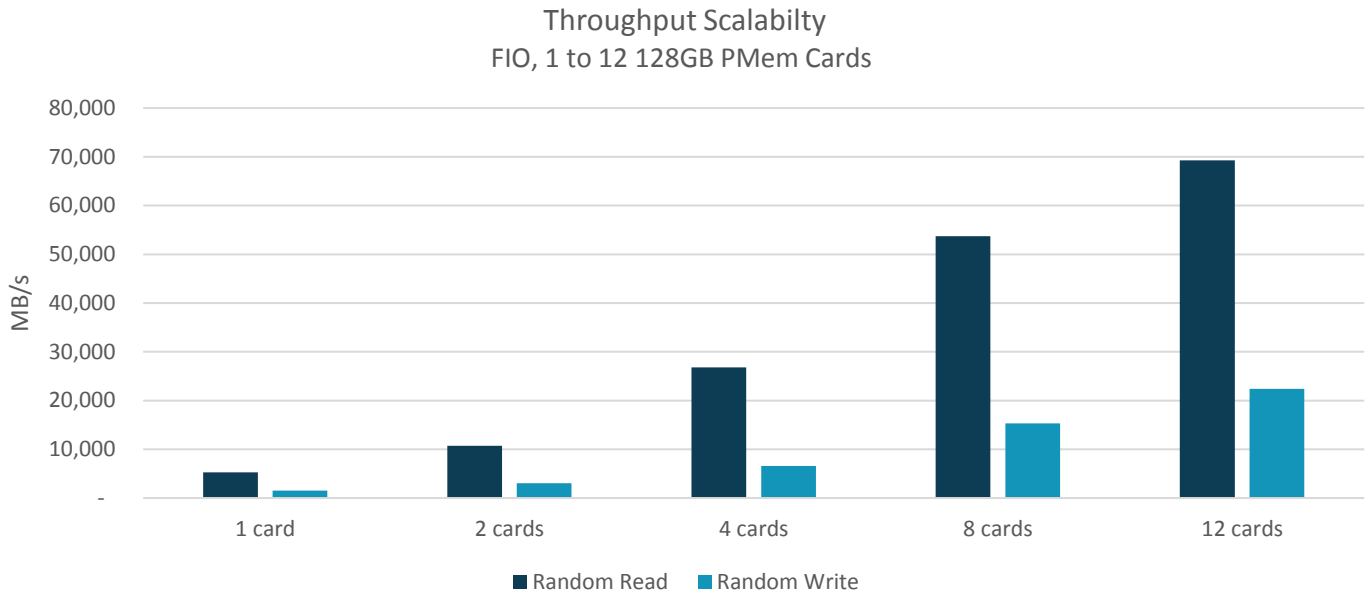
Next, we looked at the ability of the system to scale performance as PMem capacity increased. Tested configurations included one, two, four, eight, and twelve 128GB PMem cards, configured with interleaving. With multiple PMem cards configured on both sockets, VxRail created and assigned four NUMA nodes per socket. Figures 7 and 8 show near-linear scalability of random read and write IOPS and throughput as the amount of PMem capacity increases.

Figure 7. Near-linear IOPS Scalability



Source: Enterprise Strategy Group

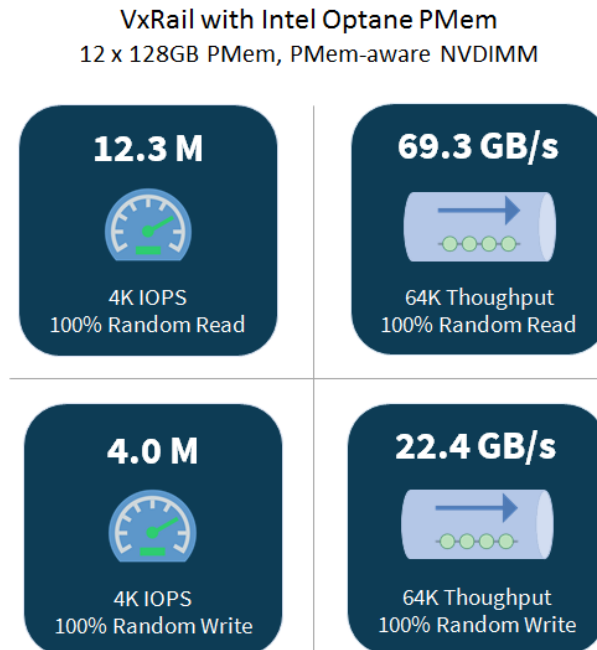
Figure 8. Near-linear Throughput Scalability



Source: Enterprise Strategy Group

Figure 9 shows the “four corners” performance using 12 x 128GB PMem cards.

Figure 9. High-performance IOPS and Throughput, 12 PMem Cards



Source: Enterprise Strategy Group

i Why This Matters

According to ESG research, the top-five most-cited benefits organizations realized by deploying HCI include improved TCO and scalability, faster deployment time, and simplified management.⁴ These key benefits, by minimizing costs and management effort, make infrastructure work harder and more efficiently for the business. However, performance challenges of consolidated virtualized workloads have prevented organizations from placing critical analytics, OLTP, and other applications on HCI.

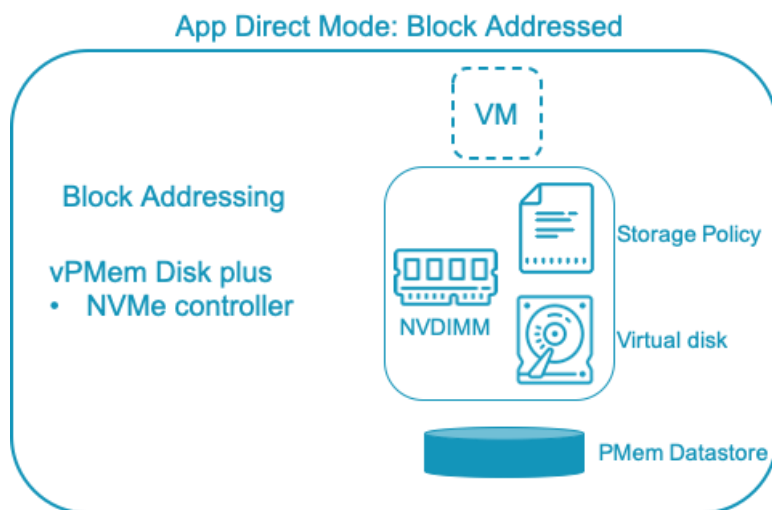
ESG validated that VxRail with Intel Optane PMem delivered more than 1 million IOPS and more than 5GB/s throughput using a single 128GB card, and near-linear scalability as PMem capacity increased. A single host with 12 x 128GB PMem cards achieved more than 12M read IOPS, more than 4M write IOPS, more than 69GB/s read throughput, and more than 22GB/s write throughput.

These results indicate that VxRail with Intel Optane PMem can not only speed performance of traditional HCI workloads such as VDI, but also support performance-intensive transactional and analytics workloads. This solution offers orders-of-magnitude faster performance than traditional storage, while including data persistence with no increase in fabric traffic. Compared to an all-flash VSAN configuration, a VxRail host with 12 x 128 GB Intel Optane PMem delivered up to 80x more read IOPS and 28x more throughput. Mission-critical databases and enterprise applications that store large amounts of data in working memory, including in-memory databases, can reside on VxRail with high performance, faster restart, and increased uptime.

Simulated Application Testing

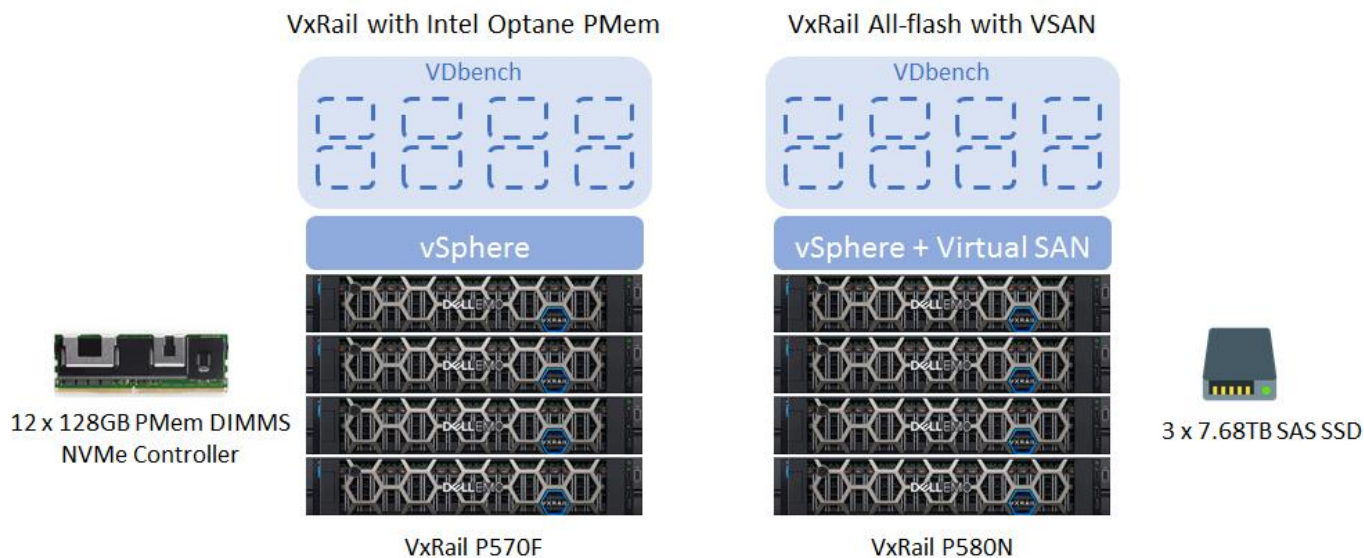
Next, ESG reviewed testing designed to simulate real-world workloads using vPMem disk. In this method, the vDisk was created on the local PMem datastore of the host and associated with a storage policy. This mode is completely transparent to applications and operating systems, requiring no changes.

We compared results of VxRail configured with vPMem disks using a virtual NVMe controller to VxRail using VSAN configured with all-flash storage. The PMem tests were run using four VxRail P570F nodes configured with Intel Xeon Gold 6254 CPUs, 12 x 32GB DRAM, and 12 x 128GB PMem; it supported two VMs per host with 8 x 25GB vPMem disk per VM. The all-flash VSAN environment was four similarly configured VxRail P580Ns with two VMs per host, each configured with 8 x 200GB vmdk. The VDBench utility was used to generate a workload.



⁴ Source: ESG Master Survey Results, [Converged and Hyperconverged Infrastructure Trends](#), October 2017.

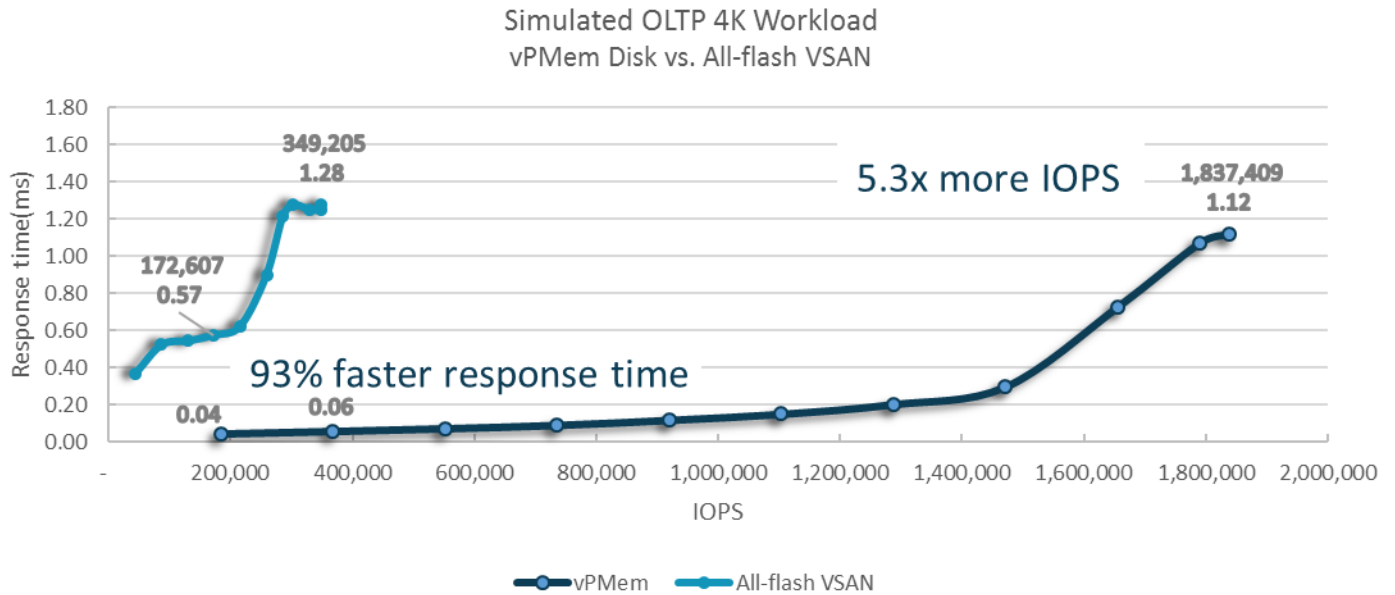
Figure 10. Simulated OLTP Workload Test Bed



Source: Enterprise Strategy Group

First, we looked at an OLTP workload consisting of 4K random I/O, 70% read/30% write. This simulates typical transactions such as retail sales, web commerce, ATM transactions, and similar tasks that involve inserting, deleting, and updating small amounts of data in a database. After an initial test run to determine the maximum IOPS and throughput, the test was run in 5-minute intervals starting at 10% of max IOPS and incrementing by 10% with each new run.

In this test, the PMem-configured VxRail delivered more than 1.8M IOPS, 5.3X more than the all-flash VxRail system. In addition, response time for the PMem VxRail remained under 0.2 ms for most of the test, while the all-flash system response time began at 0.37 ms and quickly slowed to 1.28 ms. At 10% of max IOPS, vPMem was driving 184K+ IOPS with 0.04 ms response time; VSAN achieved comparable IOPS at 40% of max, and response time was 0.57ms. The vPMem configuration was 93% faster at the beginning and averaged 72% faster response time throughout the test (see Figure 11).

Figure 11. Simulated OLTP Workload: vPMem Disk versus All-flash VSAN


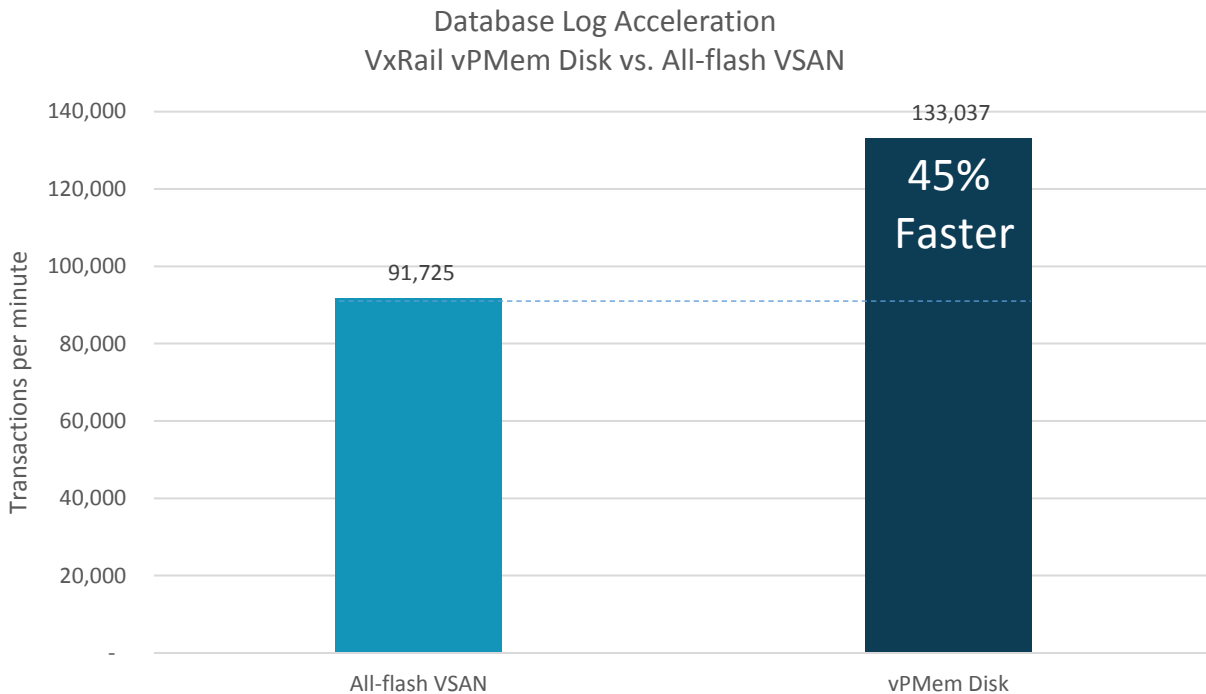
Source: Enterprise Strategy Group

Finally, we looked at results of a TPC-C-like benchmark⁵ that demonstrates database log acceleration, comparing VxRail configured with vPMem disk to a VxRail all-flash VSAN. For the VSAN configuration, all data and log files resided on VSAN; the PMem configuration stored only the 64GB log file on PMem. The test bed was four VxRail P570Fs, each with Intel Xeon Gold 6254 CPUs, 392 GB of RAM, 25Gb Mellanox networking, 745GB SSD cache, and 3 x 1.75GB SSD disks. The PMem system included 2 x 128GB cards with RHEL VMs on VMware ESXi.

MariaDB was used to create a MySQL database configured to load data into 100 data warehouses with 32 concurrent connections and an initial database size of 20 GB. The test simulated multiple transaction types and concurrent transactions of different types and complexity. It was run for one hour, tracking the number of transactions per minute. PMem VMs were built with the XFS filesystem; all VMs included 24 vCPUs and 64GB RAM.

As Figure 12 shows, the vPMem configuration provided 45% more transactions per minute. This demonstrates that for some use cases, only a small amount of PMem capacity in a host can add significant value.

⁵ "TPC-C-like" indicates that the results are not certified by TPC.org.

Figure 12. VxRail with vPMem Disk Speeds Database Log Acceleration

Why This Matters

ESG validated performance for VxRail with Intel Optane PMem used as block-addressed storage with OLTP and RDBMS workloads. The OLTP results demonstrated 5.3x more IOPS and 93% faster response time compared with an all-flash VSAN configuration; it can deliver the performance that mission-critical applications such as stock trading, web-based transactions, HPC, and gaming demand.

In addition, RDBMS testing demonstrated that with easy-to-install and cost-efficient Intel Optane PMem, IT administrators can significantly accelerate database logs. This is important because when log writes are slow, they become a choke point that drags down the entire database process, delaying transactions and analysis. The ability to accelerate that process at low cost with a small amount of persistent memory capacity, even a single card, delivers significant benefit.

The Bigger Truth

In recent years, organizations have focused on digital transformation, working hard to make optimal use of digital assets to better understand customer needs, inform strategy, and improve operations. According to ESG research, the most common objectives for digital transformation are to become more operationally efficient, and to provide better and more differentiated customer experience.⁶

To accomplish these goals, IT organizations employ a wide range of applications with varying needs. Many have realized that consolidating diverse applications on HCI can greatly improve efficiency of infrastructure and management, but only if performance demands are met. Application performance requirements have prevented many organizations from using HCI for workloads such as real-time analytics, OLTP, and high-performance computing.

The Dell EMC VxRail HCI solution now offers the option of adding Intel Optane persistent memory to increase performance without high cost. Intel Optane PMem provides a new performance tier between DRAM and SSD, with three key benefits: larger capacity than DRAM at a fraction of the cost, high-performance storage, and data persistence after power-off.

ESG validated high performance and scalability for VxRail with Intel Optane PMem in App Direct Mode,⁷ including:

- More than 1M IOPS and 5GB/s throughput with a single 128GB PMem card.
- Near-linear IOPS and throughput scalability as PMem capacity increases.
- 5.3x more IOPS and 93% faster OLTP response time than VxRail with all-flash VSAN.
- 45% faster database log acceleration than all-flash VSAN versus a single PMem card.

The results presented in this report are based on testing in a controlled environment. Due to the many variables in each production data center, it is important to perform planning and testing in your own environment to validate the viability and efficacy of any solution.

While most business applications don't need millions of IOPS and GB/s of throughput individually, when they are consolidated, the performance is divided up among all applications. VxRail with Intel Optane PMem enables organizations to consolidate more applications, even those with high-performance needs, to gain the deployment and management efficiency benefits that HCI offers. It also expands the range of VxRail-applicable workloads at the core, edge, and cloud. In addition, the ability to start with a small amount of Intel Optane PMem and grow with scalable performance fits with the HCI profile of flexibility and agility.

Most important, with VxRail with Intel Optane PMem, organizations gain more transactions, faster insight, and better business decisions while also gaining infrastructure and management cost efficiency. Enterprise-class performance that once required expensive, dedicated storage solutions is now available to any organization. For any organization looking to consolidate applications for cost efficiency while delivering high application performance, ESG recommends evaluating Dell EMC VxRail with Intel Optane PMem.

⁶ Source: ESG Master Survey Results, [2020 Technology Spending Intentions Survey](#), January 2020.

⁷ Note that an updated performance report will be available in the coming weeks for VxRail with Intel Optane PMem in Memory Mode, where it increases memory capacity at a low cost per GB.

Appendix

Table 1. Dell EMC VxRail Test Configurations

App Direct Mode Performance and Scale Testing	
Dell EMC VxRail with Intel Optane PMem	
VxRail Platform	4 * VxRail P570F, Software v. 7.0.000
CPU	Intel Xeon Gold 6254 3.1GHz* 2 (36 cores per node)
Cache per node	Samsung 800Gb NVMe
Data Storage per node	3* 7.68TB Toshiba SAS SSD
ESXi Version	VMware ESXi 7.0.0 build 15525992
VM OS	RHEL 8.1
Network	25 GbE Mellanox
Intel Optane PMem	1 *, 2 *, 4*, 8 *, 12 * 128GB Intel Optane PMem cards configured with interleaving
Test Harness	FIO version 3.7

Simulated Application Testing: OLTP		
	Dell EMC VxRail w/ Intel Optane PMem	Dell EMC VxRail All-Flash
VxRail Platform	VxRail P570F	VxRail P580N
CPU	Intel Xeon Gold 6254, 3.1GHz* 2 (36 cores per node)	
Cache per node	12 * 32GB DRAM	
Data Storage per node	3* 7.68TB Toshiba SAS SSD	
ESXi Version	VMware ESXi 7.0.0 build 15525992	
VM OS	RHEL 8.1	
VM configuration	2 VMs/host, each with 8x25GB vPMem disk	2 VMs/host, each with 8x200GB vmdk
Network	25 GbE Mellanox	
Intel Optane PMem	12 * 128GB cards, vPMem disk, NVMe controller, configured with interleaving	
Test Harness	VDbench version 5.04.07	

Simulated Application Testing: Data Warehouse Log Acceleration		
	Dell EMC VxRail w/ Intel Optane PMem	Dell EMC VxRail All-Flash
VxRail Platform	VxRail P570F	VxRail P580N
CPU	Intel Xeon Gold 6254, 3.1GHz* 2 (36 cores per node)	
RAM	392 GB	
Cache per node	745GB Toshiba PX04SMB080 SAS SSD	
Data Storage per node	3* 1.75TB Toshiba PX04SMB080 SAS SSD	
ESXi Version	VMware ESXi 7.0.0 build 15525992	
VM OS	RHEL 8.1	
VM Configuration	24 vCPUs, 64GB RAM	VSAN Storage policy FTT1
Database	MariaDB version 10.3.17	
Network	25 GbE Mellanox	
Intel Optane PMem	2 * 128GB card; vPMem disk, NVMe controller, 64GB NVMIDD, XFS File system	

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.



Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides market intelligence and actionable insight to the global IT community.

© 2020 by The Enterprise Strategy Group, Inc. All Rights Reserved.